DOCUMENT RESUME

ED 244 981 TM 840 301

AUTHOR Patrick, Allison H.; And Others
TITLE Qualitative and Quantitative Evaluation of the Early Prevention of School Failure Program for Kindergarten Students.
PUB DATE Mar 84
NOTE 53p.; Paper presented at the Annual Meeting of the American Educational Research Association (68th, New Orleans, LA, April 23-27, 1984).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Academic Failure; *Evaluation Methods; Formative Evaluation; *Handicap Identification; *Intervention; *Preschool Children; Preschool Education; Primary Education; Program Evaluation; Screening Tests; Summative Evaluation
IDENTIFIERS *Early Prevention of School Failure; Meta Evaluation

ABSTRACT

The Early Prevention of School Failure Program (EPSF) was initiated in Peotone, Illinois, in 1971. The stated purpose was to prevent school failure through early identification and remediation of developmental learning deficiencies in four- to six-year-old children. The project components were team screening, team conference, education planning, modality instruction, student's progress evaluation, and parental involvement. These components were instrumental in achieving the program's objective. This paper reports on the methods and results of four types of evaluations: meta-evaluation, ex post facto evaluation, process evaluation, and product evaluation. The tests and screening instruments utilized by the program (the Preschool Language Scale, Developmental Test of Visual Motor Integration, Peabody Picture Vocabulary Test, Draw a Person Test, Revised Motor Activity Scale, Metropolitan Readiness Test, and California Achievement Test) are described. The qualitative and quantitative evaluations of the program are summarized. (DWH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED244981

QUALITATIVE AND QUANTITATIVE EVALUATION OF THE EARLY PREVENTION OF SCHOOL FAILURE PROGRAM FOR KINDERGARTEN STUDENTS

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Allison H. Patrick
George H. Kimball, Ph.D.
John Crawford, Ph.D.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

P. J. Watson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Planning, Research, and Evaluation Department
Oklahoma City Public Schools
Oklahoma City, Oklahoma

March 1984

Prepared for Presentation at the Annual
AERA Meeting, April 1984 at New Orleans, LA

TM 840 301

Qualitative and Quantitative Evaluation
of the Early Prevention of School Failure Program
for Kindergarten Students

BACKGROUND

The Early Prevention of School Failure (EPSF) Program was initiated in 1971 at Peotone, Illinois in response to a perceived need for early identification and intervention for children with learning and developmental deficiencies. The programmatic goal of EPSF is to: "prevent school failure through early identification and remediation of developmental learning deficiencies in 4 to 6 year old children; deficiencies that would adversely affect their school performance".[1] The goal is achieved, according to program literature, by: "the successful implementation of six identified critical and observable project components" (emphasis added).[2]

These components are:

(1) Team Screening of all kindergarten students in five "modality" areas, i.e., language, auditory, visual, fine motor and gross motor. Speech, vision and hearing are also assessed.

(2) Team Conference during which observations made by the screening team during the screening process, test scores and information from the students' parents are evaluated. Based on this conference, children are placed by "need" in an appropriate modality group.

---

[1]Werner, L. Judy, J. Larsen, C. Reed, J. and Gauntt, H. Early Prevention of School Failure Training Handbook. Peotone, IL. 1982, p.2

[2]Ibid., p. 2-3

(3) Education planning is conducted by the kindergarten teacher from recommended activities from EPSF curriculum guides. These activities are designed to enhance developmental skills in the five modality areas and emphasize practice with concrete materials.

(4) Modality instruction conducted during a 20-30 minute daily modality group allows children with learning deficiencies to practice basic skills in the areas of concern, while under supervision of the kindergarten teacher. Teachers maintain records on each child as skills are practiced and mastered.

(5) Evaluation of students' progress is accomplished using three tests from the screening battery as a post test. The results of post testing with Preschool Language Scale, Visual Motor Integration and Peabody Picture Vocabulary Test are compared to results from the original screening for indications of progress.

(6) Parental involvement is encouraged in EPSF program classroom activities and in work with the child at the home.

The results of the team screening mentioned above, are designed to categorize children as "high", "medium" or "average" developmental risks and to identify strengths and weaknesses in the five "modality" areas indicated. Tests used for screening purport to identify the primary learning modality for each child. Further, the screening instruments are to be used as bases for developing individualized learning plans that the teacher follows during the year for remediation and compensation for developmental deficiencies and enhancement of strengths. According to various literature published by the project office in Peotone, post test results indicate considerable gains in modality areas which had been identified by pre tests

as developmentally deficient. EPSF was declared an Exemplary Educational Program by the National Diffusion Network in 1974. It appears in the Fall 1980 NDN Catalog Educational Programs That Work.

EPSF was introduced into the Oklahoma City Public Schools on a pilot basis during the 1979-80 school year. Four elementary schools participated in the pilot effort during which all kindergarten and first grade students in the schools were screened and assigned risk levels during October 1979 then post tested during May 1980. A total of 126 kindergarten and 77 first grade students participated in the pilot study. The Oklahoma City School Board voted on May 3, 1982, to adopt EPSF District wide, and on May 17, 1982, to institute EPSF during the 1982-83 School Year.

PURPOSE OF THIS PAPER

The purpose of the current paper is four-fold. These are:

1. To report the method and results of a meta-evaluation conducted by reviewing evaluation studies produced by EPSF program staff. During this phase, analysis techniques were applied which were different from those used by EPSF staff. The meta-evaluation component also criticises the appropriateness of evaluation metrics use by the program staff.

2. To report the methodology and results of an ex post facto evaluation of the pilot effort. This component was accomplished by comparing the achievement scores of students who participated in the pilot study with those received by comparison students matched as closely as possible on socioeconomic strata.

3. To report on the methodology and results of a process evaluation during which various qualitative data were gathered from survey instruments, classroom observations, and interviews.

4. To report on the methodology and results of a product evaluation based on a comparative analysis of California Achievement Test (CAT) scores attained by kindergarten students in 1981, 1982, and 1983. Several other comparisons were made including correlations of EPSF program pre-post gain measures with CAT scores, effects of amount of treatment related to achievement and pre-post gains, and the relationships of the District's criterion referenced reading management system with achievement.

DESCRIPTIONS OF TESTS AND SCREENING INSTRUMENTS

Preschool Language Scale (PLS): The PLS was developed by program staff at Peotone to predict school readiness by assessing auditory, visual and motor synthesis among children ages 3 through 8 years. The scale consists of five individually administered subtests which are: (1) Visual-Vocal integration -- which requires the examinee to integrate visual-auditory-vocal association, discrimination and memory before responding; (2) Vocabulary -- which provides an effective method to assess both ability and experience levels of integrated auditory responding; (3) Auditory response -- which examines auditory synthesis at its highest level by requiring remembering, association and response to complex auditory stimuli; (4) Integrative Auditory Memory -- which is designed to assess integrative Auditory Memory of discrete phonemes represented by proper or related parts of speech; and, (5) Discriminative Visual-Auditory

Memory --which examines ability to code a consistent relationship between or among graphemes and/or phonemes. Test-retest reliability is estimated to be .773 among a small, select group of kindergarten students from in and around Peotone, Illinois. No validity data are cited in the test manual or related literature. [3]

The Developmental Test of Visual-Motor Integration (VMI) was constructed to detect children with problems in visual motor integration. The VMI consists of a number (from 15 to 24) of geometric forms that are to be copied in a test booklet. The forms, arranged in order of increasing difficulty, are copied until three forms in succession are copied unsuccessfully. An individual's score is reported as mental or test age in months. Scoring procedures contain a high degree of subjectivity, especially among inexperienced scorers. Neither construct nor predictive validity for this test has been established (or at least reported). As of 1972, reliability had been established only for a small sample from a limited population in a restricted geographic area.[4]

The Peabody Picture Vocabulary Test (PPVT) is designed primarily to measure an examinee's receptive or hearing vocabulary for Standard

---

[3] Preschool Language Scale Manual, Peotone, IL.: Early Prevention of School Failure Nationally Validated Project, 1981.

[4] Chissom, B.S., Developmental Test of Visual-Motor Integration, in Buros, O.K., (Ed.), The Seventh Mental Measurements Yearbook, Highland Park, N.J.: The Gryphon Press, 1972.

American English. Scores are recorded as raw scores for conversion to standard scores or to age equivalents in months. The test does not require reading ability and responses can be gestural or verbal. Norms were established on a standardized sample of 4200 children between the ages of 2 years 6 months, 18 years 11 months, representative of the 1970 geographic, ethnic and occupational distribution of the U.S. Split-half reliabilities are reported to be from .61 to .88 depending on test form and examinee age. Alternate forms reliability is reported to be from .73 to .91. Validity estimates in the nature of correlations to other vocabulary tests are reported to be on the order of from .20-.89. Reported correlations with achievement tests vary widely depending on the achievement test and subtest, with correlation coefficients ranging from .29 to .68. It should be noted that reliability and validity coefficients were determined by the test author's review of testing literature.[5]

The Draw a Person scale (DAP): The DAP focuses on a child's accuracy of observation and on the development of conceptual thinking. The subject is told to draw a man, a woman or a person. Scoring is based on 73 items for drawing a man, and 71 items for drawing a woman. Items are scored on an all or none basis. Raw scores are coverted to standard scores. Norms were established with a standardized sample of 2975 children between the ages of 5

---

[5]Dunn, L.M. and Dunn, L.M., Peabody Picture Vocabulary Test - Revised Manual - Forms L and M, Circle Pines, MN., American Guidance Service, 1981.

and 15 years representative of the occupational distribution of the U.S. in 1950. Scoring instructions are complete and, when used carefully, interscorer reliability is on the order of .90. As of 1972, split-half reliabilities are reported to be in the .70-.80 range. Validity coefficients are based on correlations between the Goodenough Draw-a-Man, and intelligence and special aptitude tests and are reported to be "significant and most are substantial".[6] The Goodenough-Harris Drawing test is a refinement and extension of the Goodenough Draw-a-Man Test which has only 51 scorable items.

The Revised Motor Activity Scale (MAS) is used for evaluating a child's body awareness, manual dexterity and body control. Perceptual-motor skills are a part of a child's non-verbal development. These skills involve "awareness of objects" and "information through the senses" and the "ability to perform coordinated movements". The scale is individually given and evaluates balancing, rhythm, directionality, body image, fine and gross movement, bilateral activities and left-right dominance. Scores are reported by use of a point scale. The test was normed with a sample of 138 children. Ages and geographic representation of the norming sample are unknown. Neither validity nor reliability data are reported. [7,8]

---

[6]Anastasi, A., Goodenough-Harris Drawing Test, in Buros, O.K. (Ed.), The Seventh Mental Measurements Yearbook, Highland Park, N.J., The Gryphon Press, 1972.

[7]Training Handbook, Peotone, IL., Early Prevention of School Failure Nationally Validated Project, 1982.

[8]Heinger, M., Motor Activity Scale Manual, Peotone, IL., Early Prevention of School Failure Nationally Validated Project, 1979.

In addition to program related tests, two tests were used to provide a basis for estimating achievement gains. These tests were the Metropolitan Readiness Test and the California Achievement Test.

Metropolitan Readiness Test (MRT): The MRT, 1976 Edition, consists of two levels. Level 1 is designed for use with kindergarten students or beginning grade 1 students judged tc have insufficient prereading skills. The MRT is useful in assessing the diverse range of prereading skills present in kindergarten and beginning grade 1 students. The test provides to teachers information useful in classifying students. Subtests include: auditory memory, rhyming, letter recognition, visual matching, school language and listening, and quantitative language. Level 1 of the MRT was normed for a grade 1 population with a total of 7565 students, 5400 from outside large cities, in areas where the median adult level of schooling was below 12.12 years, and 2165 first grade students from the Country's largest school systems. Average student age was reported as approximately 6.5 years. Kuder-Richardson formula 20 reliabilities are estimated to be from .68 to .78 for the subtests and .92 for the composite or total score. Predictive validity for the MRT was determined by comparing fall MRT scores with Metropolitan Achievement Test (MAT), Primary 1 scores achieved in the Spring by 1272 of the norming sample. The overall or composite MRT score correlated .61 with the MAT total

reading score and .64 with the MAT total math score.[9,10,11]

The California Achievement Test (CAT) is a test battery designed to measure the achievement of students from the beginning of kindergarten through the twelfth grade. For testing first grade students Level 11, Form C, was used. CAT Level 11 contains six subtests including phonic analysis, reading vocabulary and reading comprehension in the reading area; mathematics computation, and math concepts and applications in the mathematics area; and, a language expression area. Each of these areas yields a score indicative of ability levels. While no criterion or construct validity estimates are provided in the CAT literature, the publishers explain in detail the methodology used in constructing and trying-out the test to insure content validity. Test-retest reliability estimates for Level 11 range from .35 to .84 for the subtest areas. Total Battery reliability is reported to be .84. The CAT, Level 11, was standardized, or normed, during the Fall-Spring 76-77 period using a sample of 17,736 first grade students from various regions and backgrounds.

---

[9]Nurss, J.R., and McGauvran, M.E., Metropolitan Readiness Test Teachers Manual Part 1: Directions for Administering, Level I, Form P, New York, The Psychological Corporation, 1976.

[10]Norms on Form P of Level 1 for Special Grade 1 Population, (Metropolitan Readiness Tests Research Report Number 6), New York, The Psychological Corporation, 1977.

[11]Dykstra, R., The Metropolitan Readiness Test, The Buros, O.K. (Ed.), The Seventh Mental Measurements Yearbook, Highland Park, N.J., The Gryphon Press, 1972.

For testing kindergarten students Level 10, Form C, was used. CAT Level 10 contains seven subtests including listening for information, letter forms, letter names, letter sounds, visual discrimination and sound matching in the prereading area; and a mathematics concepts test in the math area. Each of these areas yields a score indicative of ability levels. While no criterion or construct validity estimates are provided in the CAT literature, the publishers explain in detail the methodology used in constructing and trying-out the test to insure content validity. Short term interval (2-3 week) alternate form reliability estimates for Level 10 range from .64 to .91 for the subtest areas. Total Battery reliability is reported to be .93. The CAT, Level 10, was standardized, or normed, during the Fall-Spring 76-77 period using a sample of 13,201 kindergarten students from various regions and backgrounds across the U.S.[12]

---

[12] California Achievement Test (Technical Bulletin 1), Monterey, CA., CTB/McGraw-Hill, 1979.

META-EVALUATION OF EARLY PREVENTION OF SCHOOL FAILURE EVALUATIONS COMPILED BY PROGRAM STAFF

The purpose of the part of the study is to examine critically evaluation information distributed by the Early Prevention of School Failure (EPSF) project office in Peotone, Illinois. The object is to determine the adequacy of evaluations conducted by the project office and the accuracy of subsequent claims of program effectiveness. The study is based on an analysis of evaluation results presented in various reports distributed by the project office, Early Prevention of School Failure Nationally Validated Project in Peotone. The primary data source was the Evaluation Summary, 1981. It should be noted that the program orginated references cited are not exhaustive. Other evaluation reports had been published but were not available to the current author.

Examination of available evaluation reports distributed by the project office revealed that the commonly utilized measure of success is a metric defined in terms of "months-gained-per-months-in-program". If a student had demonstrated a metric score greater than one, that student's progress was taken as evidence of program effectiveness. It should be noted that gain scores were reported only for students who had been found at "high" risk or at "moderate" risk by the program recommended pretests, i.e., relatively low scorers. There was little or no mention of gain scores expected from average students. Nor, typically, were students enrolled in an EPSF program compared

with students who were not enrolled, or, to a control group.

There is one major exception to the pattern of non comparative results, and that is a "long range project"[13] in which an unspecified number of students were selected to participate in a longitudinal study of program effectiveness. Part of these students received the EPSF "treatment", and remainder acted as "control". Both groups were tested at the beginning and at the close of the school year and again one year later with the same instruments, i.e., PPVT and VMI. Other than this project, indicators of program success appear to be based on single group, pre-post test design.

Discussion

For several reasons such as regression of scores toward the mean, maturation and socialization, the use of gain scores in a test age metric to demonstrate program effectiveness is questionable when the score is considered in a non-comparative manner.[14] There are several plausible alternative explanations to an increase in test age. Score gains greater than one can be explained and attributed to a regression effect, maturation or socialization as well as to program related or caused growth.

13Strand, K. and Werner, L. Evaluation Summary of the Early Prevention of School Failure Project A Nationally Validated Program, 1971-1981. Peotone, IL: 1981.

14Heimendinger, J. and Laird, N., Growth Changes Measuring the Effect of an Intervention, Evaluation Review: 7, 1, Feb. 1983, pp 80-95.

Examination of the project office's long range project in which a control or non-treatment group was designated, provides a vehicle with which to examine more closely reported program effectiveness.

In this project staff report, when pre- and post-test scores for the PPVT and the VMI were tested by ANOVAs, there were no significant interaction effects found between group membership (i.e., treatment or no treatment) and time in treatment. When the data points were plotted (See Figures 1 and 2) it was found that visually as predicted by the lack of an interaction effect, the slopes for the two groups were approximately parallel. Apparently the treatment group was selected with a higher initial pretest score and remained higher throughout the testing cycle. Insufficient information was available to test for differences while controlling for pretest scores. It appears that the rate of learning as defined by the test scores over time was the same for the control group and the treatment group. This seems to indicate a lack of treatment or EPSF program effect.

A major area of concern which arose during review of project documents, was the lack of any validation effort for the program developed testing instruments. In the Preschool Language Scale (PLS) Manual, program staff report a validation study in which 87 students from central Illinois participated. But the report contains only results of test-retest reliability correlations. The validity of the PLS is left open to question. What the authors report as "content" validity seems rather to be face validity, i.e., the

appearance of validity. No construct or criterion validity effort is reported. This is particularly important because experience within OCPS has indicated that the PLS is the most heavily weighted test in the determination of relative risk level or strength level. Of the program developed instruments used to determine risk level, the best that can be said is that their ability to measure the constructs they claim to measure is questionable. Further, reliability information presented in the test manuals is suspect because of the small and geographically limited sample used in reliability studies.

EVALUATION OF THE PILOT PROJECT:
"EARLY PREVENTION OF SCHOOL FAILURE"
FOR OKLAHOMA CITY PUBLIC SCHOOLS

This part of the study was an ex post facto examination of the EPSF pilot program conducted in OCPS during the 1979-80 school year, and was an attempt to identify any potential program impact on participating students' achievement test scores.

METHOD

In the 1979-80 school year, four elementary schools in Oklahoma City implemented the EPSF program on a pilot basis. Program related pretesting or screening, was accomplished for virtually all kindergarten and first grade students in the participating schools. Screening consisted of five program recommended tests administered by a screening team consisting of a teacher, a counselor, a speech pathologist and a psychometrist. These tests included the Preschool Language Scale (PLS), Developmental test of Visual Motor Integration (VMI), Peabody Picture Vocabulary Test (PPVT), Draw-a-Person Test (DAP) and the Motor Activity Scale (MAS). All first graders in the district had been tested with the Metropolitan Readiness Test, Level 1 Form P, either during the Spring of their kindergarten year or during the early part of their first grade year. All students participating in the pilot program were post-tested at the end of the year with the PLS, PPVT, and VMI. First grade students were also tested district-wide at the end of the school year with the California Achievement Test, Version C, Level 11.

At the time of EPSF screening for the pilot study, no students were identified and tested in a similar group of schools to act as a control group for comparison. Therefore, only those students participating in the pilot study were tested with the program recommended tests. The district wide achievement testing program provided the opportunity to obtain some achievement oriented pre-treatment test scores and post-treatment test scores which were comparable across the district. In order to compare the treated EPSF Pilot study group of students with non treated students, nine comparison schools were selected. Selected schools were in the general vicinity of each pilot school and had attendance areas similar to those of the pilot schools. Achievement test scores were collected for pilot study student cohorts in these comparison schools as well for grade cohorts who were currently in the pilot schools, but who had not participated in the EPSF program. These scores were used for comparison to the achievement test scores of students participating in EPSF. Comparison schools were selected to match socioeconomic strata in pilot schools in so far as was reasonable based on 1978-79 data received from the Oklahoma City Planning Department. Socioeconomic status information was derived from factor analytic study of census tracts within Oklahoma City which constituted each school's attendance area.[15] Using six variables including income, one-person households, households with children, retired household heads, professional and

---

15Davis, S. Methodology for Calculating the Socioeconomic Index Used in Long Range Planning Research and Evaluation Unit, Department of Technical Services, Oklahoma City Public Schools, undated.

managerial persons and low-income households, a factor score was computed for each census tract. These scores were divided into quartiles and were ranked as high, moderately high, moderately low and low. These quartile scores were then compared to the scores of the pilot school and the closest matches selected.

The decision to use the two achievement tests produced an additional problem in that kindergarten students participating in the EPSF pilot study had no achievement oriented pre-measure. Therefore, for comparative results kindergarten students were deleted from further consideration. While this greatly reduced the sample size, and had severe implications for the ability to generalize from this sample it was assumed that certain conclusions concerning program effects could be drawn. The number of students in EPSF first grades with useable achievement test scores was 29. The size of the group selected for comparison was 406. Of this number, 212 had both achievement test scores (CAT and MRT) and constituted the useable comparison group.

Analysis of the scores from program tests and achievement measures produced five observations (details of the analyses are available from the author of the current paper). These observations were:

1. Students in the pilot EPSF program did demonstrate gains on program measures consistant with those reported by program staff; however these gains could have been the result of maturation.

2. Students in the EPSF pilot program did tend to close the gap between high risk students and average risk students, as measured by program related tests; however, this closing of the gap may be related to a regression to the mean effect.

3. Overall, students in the EPSF pilot program scored slightly higher than did students in comparison schools on a pre test of achievement level.

4. Overall students in the EPSF pilot program consistantly scored at the same level as comparison school students on post test achievement measures.

5. Students classified as moderate or high risk scored no lower than did average students on the post test achievement measure, and high risk students scored significantly higher on the achievement pre-measure than did average students.

PROCESS EVALUATION OF
EARLY PREVENTION OF SCHOOL FAILURE

METHOD

EPSF was implemented in 68 schools in the Oklahoma City School District (OCPS) in September 1982. Extensive EPSF inservice training had been conducted for staff and teachers during Summer and early Fall 1982, in preparation for program adoption. Program related pretesting or screening, was accomplished for all kindergarten students who were enrolled and present from September 7 through October 1, 1982. Screening consisted of the five program recommended tests administered by a screening team made up of a kindergarten teacher, counselor, speech pathologist and psychometrist. The tests used included the Preschool Language Scale (PLS), Developmental Test of Visual Motor Integration (VMI), Peabody Picture Vocabulary Test (PPVT), Draw-a-Person Test (DAP) and a Motor Activity Scale (MAS), all of which were described previously. During the screening process, the screening team conducted conferences to identify modality strengths and weaknesses for each child, based on test score results. Evaluation of the screening process was reported in a separate study and is available upon request from the current author.

All students screened and found to be at a "moderate" or "high" risk for future failure were post tested during latter part of April 1983. In addition to the program screening and post testing, all kindergarten students were tested, using the California Achievement Battery, Level 10, Form C, from April 5 to 7, 1983.

Based on the results of the EPSF Program related screening and staffing by the screening team members, a level of development was determined in each of the five modality areas for each student. If a student's test scores indicated two or more modality areas which were two or more years below a student's calendar age, that student was found to be at "high risk" for future school failure. A finding of two or more modality areas at one year below calendar age resulted in a categorization of that student as a "moderate risk" for future school failure. Students who did not fall into either of these categories were classified as "average risks". Assignment of students to risk levels occurred at all schools. The numbers of students at each risk level are listed in Table 1.

TABLE 1

Numbers of Kindergarten Students
By Risk Level

RISK LEVEL	NUMBER OF STUDENTS
Average Risk	1088
Moderate Risk	1865
High Risk	504
TOTAL	3457

PROCEDURES AND RESULTS

Process Evaluation

The process portion of this evaluation was undertaken in an effort to determine the extent to which OCPS had implemented the EPSF program as it was designed by the project office in Peotone, IL. The process

evaluation consisted of a survey of selected parents of kindergarten students, interview and observational information from a group of randomly select kindergarten classrooms and teachers, and a survey of all kindergarten teachers. Each part will be addressed in the order which it occurred.

Survey of the Parents

An important part of the EPSF program, indeed a critical and observable component, is: "Parents are encouraged to become knowledgeable about the program, to volunteer in the classroom and/or work with their child at home".[16] Therefore, program knowledge among parents and willingness of the parent either to work as a volunteer in the classroom or to work at home with program materials with their child, would appear to be good indicators of program implementation.

Parents to be surveyed were selected on a stratified random basis, based on their association with a particular school, elementary or K-5/fifth year center. The strata for each of these categories were formed based on the percent of "high risk" students identified as a result of pretest data attending each school.

Three strata were formed by selecting percentage break points which yielded strata of approximately equal size. The content of

---

[16]Werner et al. 1982, p3.

of the strata were determined by a low number of high risk students, a medium number of high risk students or a high number of high risk risk students. Break points were: 0-8 percent: low; 8.1-18 percent: medium; and, 18.1 percent and above: high. This strata selection procedure yielded 19 elementary schools in the "low" stratum, 18 in the "medium" stratum and 18 in the "high" stratum. The same stratification procedure was followed for K-5/fifth year centers, but produced a less balanced structure. K-5/fifth year center stratification yielded 2 schools in the low stratum, 4 in the medium and 7 in the high. Three schools were selected by using a table of random numbers, from each of the three elementary school strata. One per stratum was chosen for the K-5/fifth year center schools. Survey instruments marked with a random number for accounting purposes only, were sent to the chosen schools under a letter to the school principal requesting cooperation in distributing the surveys by sending them home with kindergarten students. Surveys were accompanied by a stamped addressed return envelope with allowed parents to return the completed surveys anonymously via U.S. mail. A total of 800 surveys were distributed; however this included a double distribution to one school because some returned surveys from that school were found to have doubtful authenticity. The questionable surveys were discarded and a second mailing for that particular school was sent directly to parents with the request that if they had already completed one survey to ignore the second. Therefore effectively, 710 surveys were distributed for a total of 702 kindergarten students

(as of the day of distribution), 175 or 24.9 percent were returned.

With a 24.9 percent return, it can reasonably be said that a response bias exists among the results. It would be inappropriate to attempt to generalize the results of this survey to the entire District; however, the results, when taken in the appropriate context, can provide a certain insight into a certain category of parents, probably those who could be termed "most interested". A summary of the major points are listed below.

Parent Survey Summary

A. Of the parents responding, fifty percent indicated that their school had held no meeting (8.0%) or that they didn't know if their school had held a meeting (42.5%) to explain EPSF. 35.6% indicated that a meeting was held and they had attended.

B. When asked whether they had attended a parent-teacher conference with their child's kindergarten teacher, 75.4 percent responded affirmatively.

C. When asked if they knew what EPSF is supposed to do, 58.0% of the parents responded affirmatively.

D. Parents, when asked if they had been told the results of the EPSF screening, indicated that 43.5% had not been told the results and 12.4% were told the results but not the child's level of risk.

E. When asked how they had been told the results of EPSF screening 95.6% of those answering this question indicated they were told during a parent-teacher conference.

F. When asked if they had been invited to visit their child's kindergarten classroom, 51.8% indicated they had been invited and that they had gone. An additional 17.6% indicated that they had been invited but could not attend.

G. Of those responding 76.9% indicated that they had not been asked to help with EPSF in the child's classroom.

H. When asked if they would be willing to help with EPSF during classtime, 45.5% of those responding said "yes".

I. The survey asked parents whether they had recieved any materials with suggested activities from their school. Of those responding to this item, 49.1% answered affirmatively and that the materials were helpful. However 46.8% indicated that none had been received (35.1%) or, that they did not know (11.7%).

J. When asked to compare this year's kindergarten with Oklahoma City kindergarten attended by an older child, 53.0% indicated that this year's program was "somewhat better" (30.9%) or "much better" (22.1%). 25.0% rated this and previous years' programs as about the same.

K. When asked what they liked best about EPSF, 55.4% provided no response (40.0%) or indicated that they didn't understand the program sufficiently to comment (15.4%). Sixteen percent indicated that it helped parents to understand their child's weaknesses and capabilities.

L. When asked what one thing they would change about EPSF, 65.1% provided no response, 12.0% responded "nothing" or "none" and 8.6% indicated that they didn't know.

Interview and Observation Summary

Two critical and observable program components are directly affected by the action of the teacher. These are:

> "Educational Planning - Teacher selects appropriate activities recommended in the Early Prevention of School Failure curriculum guides. These activities are indexed to the developmental skills in each of the five modality areas and emphasize practice with concrete materials"; and:
>
> "Modality Instruction - During 20-30 minute daily modality groups, identified children practice the basic skills in their concern areas under the supervision of the classroom teacher, specialist, or parent volunteer. The management system provides a format for simplified record keeping as children practice and master the skills".[17]

---

[17]Werner et al, 1982, p2-3.

Obviously, the teacher's classroom behavior and her willingness to employ EPSF techniques and curricula directly impact on the success of program implementation. Because of these factors, it was deemed necessary for the evaluator to become acquainted with EPSF as it was being used in the OCPS kindergartens. The best vehicle for this was determined to be a combination interview-observation to gather information and perceptions from principals and teachers and to observe the way in which modality groups were used in the classroom. Time and resource constraints dictated that only ten schools could be visited.

Schools to be visited were selected using a random selection procedure, with one exception. Selection procedures were designed to insure in so far as possible, that schools with varying degrees of program implementation were visited. To accomplish this, the District Kindergarten Coordinator was asked to give her best estimate the degree of program implementation in each school. The Kindergarten Coordinator indicated that she did not have certain knowledge of degree of program implementation but provided estimates for a majority of elementary schools. Using these estimates 2 schools were chosen from the above average category, 2 were selected from the average category, 1 was selected from the below average category and 1 was selected by the evaluator at random from the District listing of schools. In addition to these 6 schools, 1 K-5 center and 2 fifth year centers were selected at random from the District listing of schools. This procedure was followed because the rankings for these categories of schools were all very similar. The Kindergarten Coordinator was also asked to provide the name of a teacher who had exemplary

EPSF program implementation. This classroom was scheduled to be the first visited in order to provide a basis for judging program implementation.

Prior to visiting the first or "exemplary" site, the evaluator constructed two structured interviews, one for the principal and one for the teacher, to gather information concerning program implementation. However, it became apparent after several site visits that the structured format was neither precise nor sufficient to gather the required information. Therefore, the structured interview form became a note gathering medium for unstructured interviews.

Summary of Interviews with Principals

During interview with principals it became apparent why a majority of parents had the misconception that no EPSF meeting had been conducted or had no knowledge of such a meeting. The principals were in agreement that meetings had been held early in the school year, but generally parental attendance was classified as "poor". One principal who rated attendance as "good" in comparison to other events, indicated that 15 to 20 parents (of kindergarten students) attended. When compared to the kindergarten student population of that school and making the perhaps generous assumption that only one parent for each child attended, less than 25% of the parents of kindergarten students attended the EPSF orientation.

Summary of Interviews with Teachers and Observations

Interviews with teachers combined with classroom observation very quickly revealed varying degrees of program implementation. However,

even in the classroom identified as the example for the District, the program did not appear to be implemented fully. In several other classrooms visited, the EPSF program appeared not to be implemented at all. In several instances modality groups were not being used, nor was modality instruction, as such, occurring. More will be said concerning these particular classrooms later in this paper.

Survey of the Teachers

An end of year survey was conducted among all kindergarten teachers to verify information and trends identified during the interview process and to provide yet another estimator of program implementation. Survey instruments were sent individually, via school mail to each teacher along with an addressed return envelope. A total of 101 surveys were dispatched and 74 were returned for a 73.3% return rate.

DISCUSSION

Information gathered during the conduct of the process evaluation provided a multi-faceted view of program implementation, from which one can draw several conclusions.

As indicated previously, EPSF Program literature defines certain components which are identified as critical and which must be successfully implemented for successful achievement of program goals. The foregoing evaluation specifically examined implementation of project components 3, 4, and 6. These are:

"(3) EDUCATIONAL PLANNING - Teacher selects appropriate

activities recommended in the Early Prevention of School Failure curriculum guides. These activities are indexed to the developmental skills in each of the five modality areas and emphasize practice with concrete materials.

(4) MODALITY INSTRUCTION - During 20-30 minute daily modality groups, identified children practice the basic skills in their concern areas under the supervision of the classroom teacher, specialist, or parent volunteer.

The management system provides a format for simplified record keeping as children practice and master the skills.

(6) PARENT INVOLVEMENT - Parents are encouraged to become knowledgeable about the program, to volunteer in the classroom and/or work with their child at home.[18]

When parents of kindergarten students from randomly selected schools were surveyed concerning their depth of knowledge and willingness to participate in EPSF classroom activities, a surprisingly low number responded. Even the parents from one school who received a double mailing of the survey responded with a return rate of 46.7%. The overall return rate of 24.9% limits generalizability of the results because almost certainly a significant response bias is present. However, it is reasonable to make certain conclusions concerning the general parent population if, indeed, the returned surveys are indicative of opinions and perceptions of parents who returned the survey. Certainly, a large percentage of parents did not attend an EPSF orientation meeting and based their program knowledge on the results of parent-teacher conferences.

---

[18]Werner et al, 1982, p2-3

While this is not undesireable it does indicate that the more formal and expected channel for information flow was not functioning adequately. Further, typically parent-teacher conferences occur well into the school year. Parents did not become informed of program purposes and parents' responsibility until late in the first semester. Even though the majority of parents indicated that they did know the purposes of ESPF, the large number responding negatively to the item concerning EPSF materials in the home and some unsolicited written comments tended to indicate that a number of parents believed that EPSF program consisted entirely of only the screening.

An interesting dichotomy existed between information from the parents and information received from the teachers and principals during interviews. Typically teachers indicated during the interviews, that they had made several futile attempts to get parents to volunteer to assist in the classroom. Yet in the parent survey, the group of parents responding seemed to indicate that they had not been asked to assist, and a good percentage (45.5%) indicated that they would be willing to assist during class time. One could conclude that the parents are not receiving or at least not decoding the teachers' requests for assistance.

The interview and observation sequence produces the inescapeable conclusion that EPSF as a program was not fully implemented within the district. However, degree of implementation ran a nearly complete spectrum from "nearly implemented" to "not at all implemented". In no school was a full, complete implementation observed. But, there is some reasonable doubt as to whether EPSF can be fully implemented

without regular, consistent assistance for the teacher. In the classroom identified as a standard against which to judge others, the teacher obviously had invested a great many hours of her own time to produce activities for use in the modality centers in her classroom. Further, this teacher was very conscientious about maintaining individual records and scheduling children through the modality centers. However, the teacher admitted freely that she simply did not have time to work with the high and moderate risk children on a one-to-one or small group basis. No volunteers or aides were available and so children were allowed to work on their own at the modality centers with activities that the teacher had readily available. However, in several instances, children who were identified as high risk students did not know what to do with the activities and so made up their own games which may or may not have provided the modality practice needed.

At the other end of the spectrum were several classrooms in which the teachers simply were not using the EPSF curriculum at all. This occurred for several reasons such as a new teacher not familar with EPSF or the teacher perceiving that EPSF offered nothing over and above what was done on a regular basis before adoption of the program. In fact, when these classrooms were observed, the activities were very similar to the sorts of activities recommended by the EPSF curriculum.

All the teachers observed seemed to be effective with their children whether the EPSF program was being used or not. More will be said concerning those schools in which EPSF was not being used in the Product Evaluation portion of this paper.

There existed a major flaw with the instrument used to survey teachers in that no definitions were given for categories used to describe frequency, i.e., frequently, occasionally, or rarely. Therefore, it is conceivable, indeed probable that there is overlap among juxtaposed categories. But, information provided by the survey completed by the teachers seemed to support and confirm the conclusions resulting from the observations and the interviews. However, some interesting anomalies arose which can be the subject of some speculation. Although a large majority (78.1%) of the teachers agreed that they knew how EPSF should be implemented and 63.0% indicated that they were able to implement EPSF, only 48.6% indicated that they frequently or more often used one-on-one or small group modality instruction for high risk students. Further, only 49.3% indicated that they used EPSF modality centers frequently or more often. Also, although parental involvement is an important part of program implementation, 60.3% of the teachers sent activities to be done at home only occasionally or less frequently.

Although there seemed to be some confusion on the part of teachers as to what constituted full implementation of EPSF, a review of the EPSF program as it was originally designed for implementation reveals that the amount of assistance and time required for full implementation may somewhat exceed district expectations.

Mrs. Werner, when outlining staff requirements to implement EPSF states that: "2. A team approach including the following specialists is essential to early diagnosis and remediation. (Emphasis added.)

Director - Responsible for integration of the program.

Social Worker - Responsible for parent contacts, referrals and behavioral modification.

Learning Disability Consultant - Responsible for identifying learning problems, providing methods and materials, itinerant remedial L.D. kindergarten services and teacher workshops.

Teacher - Responsible for L.D. kindergarten remediation and itinerant services.

Speech Therapist - Responsible for speech therapy.

School Psychologist - Responsible for testing.

3. Parent, regular teacher and administrators involvement is essential to a successful program."[19]

CONCLUSIONS

In view of the foregoing, several conclusions are reasonable. These are:

A. Even though testing is not popular because of the time it takes, teachers rather liked the information generated from the test results. Further, the closure received from looking at a pretest and a post test yields some degree of feedback to teachers about their own effectiveness.

B. Very probably parents, and in some cases teachers, were not well informed concerning the overall aspects of EPSF.

C. Parents seemed not to be involved in assisting their children with EPSF related at home activities, not did teachers make regular use of at home activities from program materials.

---

[19]Werner, L. and Watson, G. Early Recognition of Children with Learning Disabilities. Illincis School Research, 10, 1, Fall 1973, p44.

D. Program implementation across the District varied to a large degree; however even in exemplary classrooms, EPSF was not fully implemented as designed. Further, it may not be possible to implement the program fully without extensive use of parent volunteers or other resources such as aides.

PRODUCT EVALUATION
EARLY PREVENTION OF SCHOOL FAILURE

INTRODUCTION

Although the process evaluation tended to indicate that the EPSF program was not fully implemented within the Oklahoma City Public School District, it appeared to be partially implemented and indeed some teachers were making a valiant attempt to implement the program fully. This fact is underscored by the number of minutes of EPSF treatment per week reported by most teachers. A product evaluation, therefore, should not detect gains as large as those occurring if the program were fully implemented but partial implementation it was assumed would have at least some positive effect on gain scores. Those classrooms in which the program was not implemented should produce gain scores different from those classrooms in which the teacher was making a full effort to implement EPSF. This portion of the evaluation will therefore discuss the impact of EPSF on district wide gain scores, differential gain scores of implementing classrooms vs. those not implementing and the effect, if any, of partial implementation on achievement scores.

Gain Score Analysis

Typically, evaluations conducted by EPSF program staff rely on an analysis of gain scores as demonstrated by the results of pre- and post tests for indications of program effectiveness. This portion of the product evaluation of EPSF will parallel those studies.

During the period from September 7 to October 1, 1982 every kindergarten student present within the OCPS was screened in accordance with previously outlined EPSF program procedures. Using these measures as pretests and the results of the screening team conference procedure for selection criteria, those students identified as high or moderate risks for future school failure were post tested during early May 1983. The post testing consisted of three of the five tests used for screening. Subtracting the pre-measure score from the post-measure score yields gain scores for each test. The time in EPSF program, although varying is considered to average 8 months for each student. Listings of the mean test scores, the number of students taking each test and the average gain score for each test pair are in Table 3.

TABLE 3

Average Test and Gain Scores

Test	Number of Students Participating	Mean Score
PLS Pretest	3430	67.5 (Months)
PPVT Pretest	3366	54.4 (Months)
VMI Pretest	3441	59.7 (Months)
PLS Post test	2248	88.0 (Months)
PPVT Post test	2238	61.2 (Months)
VMI Post test	2269	69.2 (Months)

Gains as Measured By:	Number of Students Participating	Mean Gain Score	Months Gained/ Months in Program
PLS	2067*	26.2 (Months)	3.3
PPVT	2053*	10.5 (Months)	1.3
VMI	2090*	11.8 (Months)	1.5

*NOTE: A student must have scores for both pretest and post test recorded to have a gain score.

It would appear to be reasonable, that if the program were at least partially effective, one could expect to see a differential effect among the risk levels. That is, one would expect students who were found to be at high risk to gain more than would students who were found to be at moderate or average risk. In fact, this is precisely what occurred. The gain scores by test and by risk level are indicated in Table 4.

TABLE 4

Gain Scores by Risk Level and Test

Test	Risk Level	Gain Score	N	Months Gained/ Months in Program
PLS	Average[1]	21.2	102	2.6
	Moderate	25.1[2]	1574	3.1
	High	31.9[3]	391	4.0
PPVT	Average	9.3	98	1.2
	Moderate	11.1[4]	1523	1.4
	High	9.2	404	1.2
VMI	Average	9.0	108	1.1
	Moderate	12.0	1569	1.5
	High	11.6[5]	411	1.4

1. Although not originally scheduled to be post tested, 108 average risk students were included in the post tested group.

2. Difference from average and high risk is significant at the .05 level.

3. Difference from average and moderate risk is significatn at the .05 level.

4. Difference from high risk is significant at the .05 level.

5. Difference from average risk is significant at the .05 level. (N.B. differences were determined using a non-orthogonal ANOVA and harmonic means for Dunn's Multiple Comparison test.)

Program effectiveness could also be indicated by a positive relationship between the amount of time spent in treatment and the amount of gain. In this case, the amount of time in treatment is a self-report by teachers of the amount of time in minutes they spent per week using EPSF techniques in their classrooms. The minutes of treatment variable was correlated with gain scores, results are in Table 5.

TABLE 5

Correlation of Minutes of Treatment Per Week with Gain Score

Gainscore	N	Correlation Coefficient	P-Value
PLS	1955	.1103	.0001
PPVT	1910	-.015	.51 (N.S.)
VMI	1963	-.131	.0001

The information in Table 5 clearly shows that the minutes of EPSF treatment per week and gain as measured by the PLS have a relationship which is positive and statistically significant. It should be pointed out that a correlation coefficient of .11 is considered a weak correlation. Significance arises probably because of the large number of scores being correlated (N = 1955). Table 5 also shows that there is essentially no relationship between gain as measured by the PPVT and minutes of treatment. It further shows a weak but significant negative relationship between amount of treatment and gain measured by the VMI.

Differential Effects of Degree of Implementation

As was indicated above, there is some question of full program implementation within the district; however, during the interview observation period, three schools were identified in which EPSF was not implemented in any degree. There were several reasons for non-implementation, but it is intuitively obvious that if program implementation causes positive gain scores then non-implementation should produce gain scores which are different, especially different from the school in which the program was more fully implemented, i.e., the school nominated by the District Kindergarten Coordinator as the exemplary EPSF school. However, when the data were analyzed to determine if there were differences in mean gain scores among the schools, it was found that, while some differences did exist among various schools, there was no difference between any of the non-implementing schools and the exemplary school when considering gain as measured by the most sensitive instrument, the PLS and using Tukey's Studentized Range (HSD) test. While this sample is very small, even to the point of being anecdotal, the results do point out that large gains can be achieved even in the absence of a 'treatment' effect.

Achievement Scores

When considering the effect of EPSF implementation on achievement scores it becomes necessary to consider achievement scores district wide for kindergarten students as measured by the CAT Level 10, Form C. However, the achievement score for 1983 is meaningless taken by itself and so must be compared with the only reasonable population of kindergarten students available, that is achievement as measured by the results from the previous year's CAT tests. Because of data availability the comparison only includes 1981, 1982, 1983 and, as such, is barely sufficient to establish a trend. Recognizing this limitation, the CAT total reading and math NCE scores for kindergarten students in 1981, 1982 and 1983 are listed in Table 6 below.

TABLE 6

California Achievement Test
Major Area Scores

	1981	1982	1983
Reading	50.46	50.77	49.25
Math	51.85	53.21	51.57

When tested to determine if the mean scores listed in Table 6 were significantly different, two non-orthogonal One Way ANOVAs were used. The summaries in Table 7 resulted.

TABLE 7

Results of One Way ANOVAs
CAT Reading Scores

Source	df	Mean Square	F-Value	P-Value
Between Years	2	2130.678	4.24	.0144 (Sig at .05)
Within Years	9802	502.167		

CAT Math Scores

Source	df	Mean Square	F-Value	P-Value
Between Years	2	2551.84	5.01	.0067 (Sig at .01)
Within Years	9802	509.80		

Recognizing that the large N yields power which may be sufficient for the ANOVA to detect trivial differences, a comparison of individual means using Dunn's Multiple Comparison Test was made. This test revealed that there were differences between the 1983 CAT reading score and the reading scores occurring in each of the other two years. In other words, the CAT total reading score for 1983 is statistically and educationally significantly lower than 1982 or 1981 ($t = 3.76$, $t = 3.26$ respectively, t (critical) $3,\infty = 2.39$). The math scores, while still lowest in 1983, are only significantly lower than those from 1982. There is no significant difference between 1983 and 1981 CAT math scores (t (1982) = 4.14, t (1981) = .72, t (critical) $3,\infty = 2.39$).

EPSF is designed to assist children who are deficient in skills required for school readiness, therefore, it was hypothesized that EPSF might assist those scoring in the lower 25 percent on the achievement measures. The numbers of students and the proportion of students falling at or below the 25th percentile for each year were determined.

The results of this determination are in Table 8.

TABLE 8

Students At or Below the 25th Percentile
on Achievement Measures

READING

Year	Number of Students Below 25%	Percent of Students
1983	997	28.2
1982	852	26.3
1981	768	25.7

MATH

Year	Number of Students Below 25%	Percent of Students
1983	1058	29.8
1982	887	27.1
1981	768	25.6

Table 8 clearly shows that in 1983 there was a greater percent of kindergarten students below the 25th percentile in reading and in math than in either of the preceding two years.

It was also deemed important to examine the relationship between minutes of EPSF treatment per week and scores on the District Achievement measures. The results of these analyses are in Table 9 on the following page.

TABLE 9

Correlation Treatment Amount with Achievement Test Scores (NCE)

CAT READING

	N	Correlation Coefficient	P-Value
Minutes of Treatment	1940	-.098	.0001

CAT MATH

Minutes of Treatment	1957	-.093	.0001

The information in Table 9 seems to indicate that there is a weak but statistically significant tendency for more treatment to be related to lower achievement scores, or less treatment to be related to higher achievement scores.

DISCUSSION

The process portion of this paper emphasizes that the EPSF program was not fully implemented in OCPS. However, at the various schools there were degrees of implementation which ranged from not-at-all to that approaching full implementation. It is impossible to determine, within the limitations of the data, an estimate of the percentage of implementation; however, it was apparent that some implementation had occurred. In view of Werner's 1973 article, it becomes apparent that EPSF was begun as a program designed specifically to meet the needs of LD and EMH kindergarten students and as such was staff-heavy. It further appears that the original EPSF

program was translated directly to all kindergarten students with little change in program activities but relying on the abilities of the teacher to implement the activities. Other major differences exist between the program as originally conceived and currently in place in OCPS. These include pre-school summer testing of kindergarten students, service for identified children by LD specialists and full day kindergarten in the original program. However, even with a program modified for use in an OCPS classroom, it is doubtful that the teacher by herself is able to implement EPSF without dependable, regular assistance to conduct modality instruction and assist in classroom activities.

In the absence of a fully operational program, it is at best tenuous to draw causal inferrences. (However, there is no evidence of process information in the evaluations accomplished by EPSF program staff. That is, although program evaluations published by the program office claim EPSF success, there is no evidence of program implementation.)

With a single group pretest - post test design we can draw only tentative and not causal conclusions. However, among those conclusions is that the gain scores achieved by students in a classroom at the implementation end of the implementation - non-implementation continum, are not significantly greater than students at the non-implementation end. The degree to which this nearly anecdotal information can be generalized is an imponderable. However, this factor does support the conclusion that increases in gain scores can and do occur in the complete absence of treatment.

This discussion of gain scores, while not totally uninteresting because that is the metric chosen by the EPSF program office to support claims of program effectiveness, may serve only to mask the real issue, that is program impact on achievement. Because of the sporadic amount of program implementation and several other factors it is inappropriate to unequivocally assign credit or blame, but that does not change the fact that in 1983 scores achieved by kindergarten students in total CAT reading were significantly lower than were scores in either 1981 or 1982. Further, math scores for 1983 were significantly lower than 1982 scores and lower, although not significantly so, than those achieved by 1981 kindergarten students. It is entirely possible that the kindergarten students of 1983 have some quality, as a class, which produces lower achievement scores. However, a comparison of post test scores achieved by students in the 1979-80 pilot test with post test scores recorded by students in 1983 revealed no obvious differences. While, strictly speaking, these groups are not comparable, this seems to support a conclusion of essentially no major differences between this group and that kindergarten class. The extent this can be generalized is, of course, unknown, and it is possible, although in the opinion of the evaluator unlikely, that the 1983 class of kindergarten students scored lower on achievement measures because of some population deficiency. Only tracking these students through several years of achievement testing will provide some interpretable information.

Another interesting result was the small but significant negative correlation between minutes of treatment and achievement scores.

This indicates a weak but statistically significant tendency for an increase in the number of minutes of treatment to be related to a decrease in CAT scores.

SUMMARY AND CONCLUSIONS

The process evaluation tended to support a conclusion that EPSF was not fully implemented within OCPS. The amount of implementation for any particular school appeared to lie along a continuum which extended from no implementation to about as much as could be done without outside assistance. There was a gap in understanding of the purposes and the objectives of EPSF by parents and this probably produced difficulty for the teacher in obtaining outside assistance from the parents, both in the form of in-school volunteers and in working with students in the home.

Teachers and administrators appeared to like the EPSF program and felt that it assists children. Teachers particularly appreciated support, in the form of test scores, for their conclusions based on observational information of student developmental inadequacy. Being able to see student progress in the form of pre- to post test comparisons provided feedback for teachers which they could use in a constructive way. Teachers liked and used the materials from the EPSF Curriculum but they disliked the program related record keeping.

Analysis of the pre- and post test scores and gain scores tended to indicate that EPSF treatment, in the form of an EPSF curriculum, was not necessary to obtain the magnitude of gain score used as an indicator of program effectiveness by the EPSF Program office.

However, because of implementation problems this conclusion is tenuous. Finally, the treatment amount, as measured in minutes per week was related in a weak negative direction to the achievement scores.

While it is questionable that the program can be implemented fully by teachers without outside assistance, there is no way to determine with the currently available information whether a fully implemented program would have a positive effect on achievement but the data seem to indicate that it would not.

FIGURE 1

PPVT TREATMENT VS. CONTROL

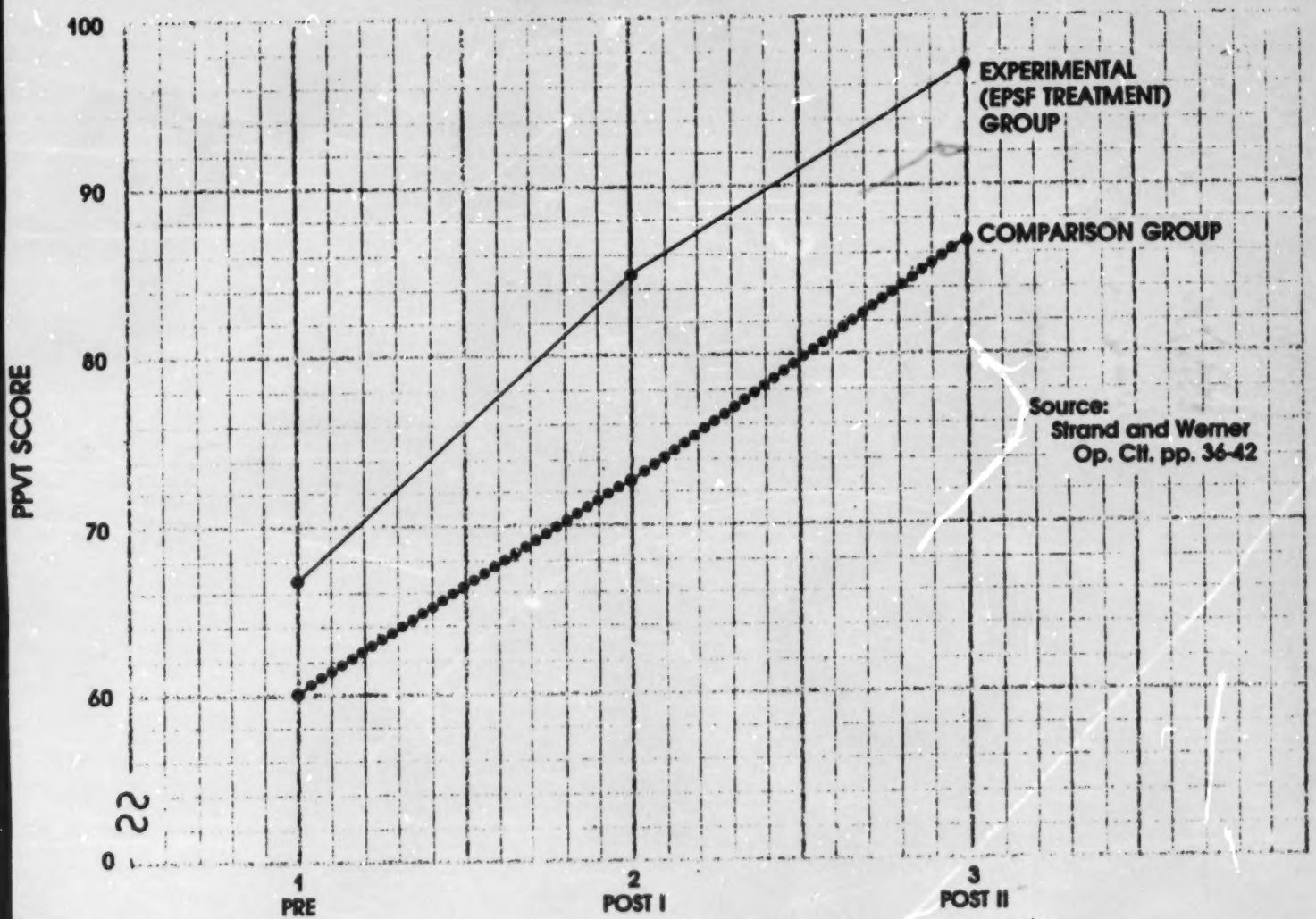
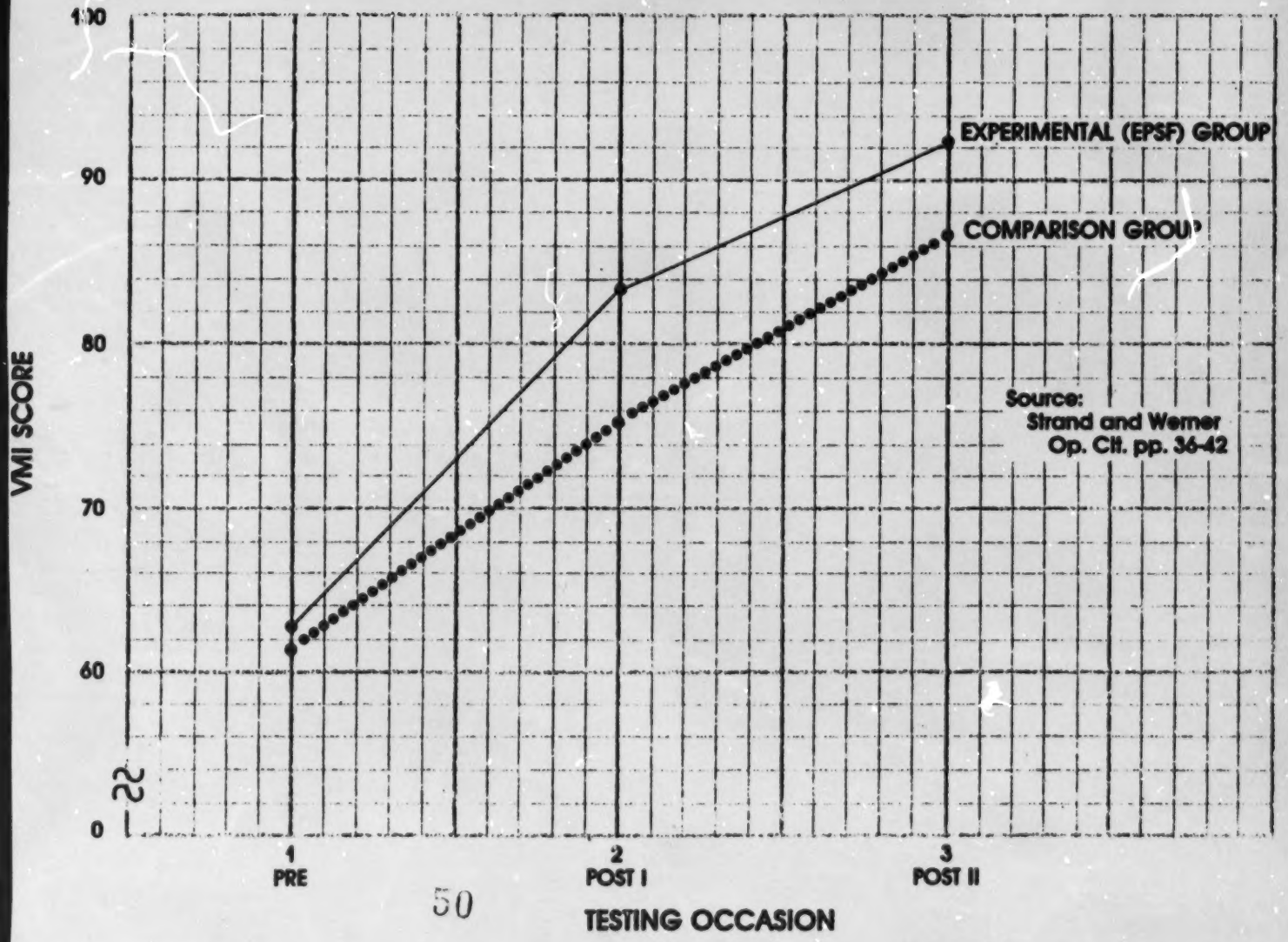


FIGURE 2

VMI TREATMENT VS. CONTROL



REFERENCES

American Psychological Association - American Education Research Association - National Council on Measurement in Education. Joint Standards for Educational and Psychological Tests, as reprinted in Buros, O.K., (Ed) Tests In Print II, Highland Park, N.J.: The Gryphon Press, 1974.

Buros, O.K., (Ed) The Seventh Mental Measurements Yearbook. Highland Park, N.J.: The Gryphon Press, 1972

California Achievement Tests Forms C and D Technical Bulletin 1. Monterey, CA: CTB/McGraw-Hill, 1979.

Campbell, D.T., and Stanley, J.C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally College Publishing Co., 1966.

Cook, D.C., and Campbell, D.T. Quasi-Experimentation Design and Analysis Issues for Field Settings. Chicago: Rand McNally College Publishing Co. 1979.

Cronbach, L.J. Evaluation for Course Improvement. In Gronlund, N.E. (Ed). Readings in Measurement and Evaluation. New York: The Macmillan Company, 1971

Cureton, E.E. Validity, Reliability and Baloney. Educational and Psychological Measurement, 1950, 10, 94-96.

Goodenough, F.L. Measurement of Intelligence by Drawings. (Extracted Copy) Publisher unknown, 1926.

Early Prevention of School Failure Nationally Validated Project Training Handbook. Peotone, Ill. 1982.

Heimendinger, J. and Laird, N. Growth Changes Measuring the Effect of an Intervation, Evaluation Review 7, 1, Feb. 1983, 80-95.

The Joint Committee on Standards for Educational Evaluation. Standards For Evaluations of Educational Programs, Projects, and Materials. New York: McGraw-Hill Book Company, 1981.

Kerlinger, F.N. Foundations of Behavioral Research. New York: Holt, Rinehart and Winston, Inc. 1973.

Messick, S. The Standard Problem Meaning and Values in Measurement and Evaluation. American Psychologist, 1975, 30, 955-966.

References cont'd.

Motor Activity Scale Manual - Early Prevention of School Failure. Peotone, Ill.: 1979

National Diffusion Network, Educational Programs That Work, Department of Education, Seventh Edition, Fall 1980

Nurss, J.R. and McGauvran, M.E. Metropolitan Readiness Test Teacher's Manual Part II: Interpretation and Use of Test Results. New York: Harcourt Brace Jovanovich, Inc. 1976

Norms Tables for Large City School Systems (Metropolitan Readiness Tests Research Report Number 3). New York: Harcourt Brace Jovanovich, Inc. 1976

Norms on Level I for End of Kindergarten (Metropolitan Readiness Tests Research Report Number 5). New York: The Psychological Corporation, 1977

Patrick, A. Interim Report Evaluation of the Early Prevention of School Failure Program for Oklahoma City Public Schools. Oklahoma City: Oklahoma City Public Schools, 1983a

Patrick, A. Evaluation of the Pilot Project: "Early Prevention of School Failure" for Oklahoma City Public Schools. unpublished paper, 1983b

Patrick, A. Meta-Evaluation Early Prevention of School Failure Evaluations Compiled by Program Staff unpublished paper, 1983c

Preschool Language Scale Manual - Early Prevention of School Failure Nationally Validated Project. Peotone, Ill.: 1981

Scriven, M. The Methodology of Evaluation. in Smith, B.O. (Ed). Perspectives of Curriculum Evaluation - AERA Monograph Series on Curriculum Evaluation 1. Chicago: Rand McNally and Company, 1967

Strand, K. and Werner, L., Early Prevention of School Failure Nationally Validated Project -- NDN Adopters (Evaluation Report for 1979-80. Peotone, Ill.: undated.

Strand, K. and Werner, L., Early Prevention of School Failure Nationally Validated Project -- NDN Adopters (Evaluation Report for 1980-81.

Strand, K. and Werner, L., Evaluation Summary of the Early Prevention of School Failure Project A Nationally Validated Program, 1971-81. Peotone, Ill.: undated.

References cont'd.

Werner, L. and Watson, G. Early Recognition of Children with Learning Disabilities. Illinois School Research, 10, 1, Fall 1973, 41-46

Williams, B., A Sampler on Sampling. New York: J. Wiley and Sons, 1978